UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

UNITED STATES OF AMERICA - v. - JOSHUA ADAM SCHULTE, Defendant.	**DECLARATION** S3 17 Cr. 548 (JMF)

PATRICK LEEDOM, pursuant to Title 28, United States Code, Section 1746, declares under penalty of perjury:

1. I am currently employed as an Incident Manager responsible for Incident Response and Digital Forensics in the Digital Security and Resilience department of the Microsoft Corporation. From 2014 to 2022, I was employed as a Lead Cyber Security Engineer at The MITRE Corporation. In that capacity, I worked exclusively as a contractor for the Federal Bureau of Investigation ("FBI"), serving as a computer scientist responsible for digital forensics and incident response in the Technical Analysis Unit of the FBI's Cyber Division. I also supported the FBI's Cyber Action Team ("CAT"), which is the FBI's on-call component of special agents and computer scientists to conduct rapid response to cyber-events by deploying to the locations of compromise around the world. I both personally deployed with CAT on more than five occasions and also personally developed several components of the forensic software used by CAT to conduct cyber-incident response. As I result of developing one of those tools, our team received the Attorney General's Award for Excellence in Information Technology. I would estimate that during my eight years working for the FBI, I supported approximately 100 investigations of various types, including criminal cases, counterintelligence cases, and hacking by nation-state

adversaries. I am a GIAC-certified digital forensic examiner, and I have taken dozens of courses and trainings in digital forensics. I hold a Bachelor of Science degree in Computer Science.

2. In my capacity as a federal government contractor assigned to work as an FBI computer scientist, I have provided assistance to the FBI special agents and Department of Justice prosecutors investigating this case. In February 2020, I testified as an expert witness in this matter. I make this affidavit based on my personal participation in this investigation and prosecution, as well as my training and experience in digital forensic analysis.

3. I have been provided with a copy of the affidavit of Steven M. Bellovin dated April 22, 2022 (the "Bellovin Affidavit"), and I make this affidavit to respond to certain claims made in the Bellovin Affidavit regarding my work in this matter and the ability of the defendant and his expert to review and dispute the conclusions that I made.

4. The version of the Bellovin Affidavit that I was provided is heavily redacted, and includes only limited information about the type of analysis that the defendant's expert wishes to conduct. I did, however, meet with the defendant's expert well in advance of the previous trial in this matter at the request of the Government, in order to explain the FBI's forensic analysis, the types of data sources that were available in the case, and to discuss what if any additional material he might require to review or analyze data further. I was willing to have additional such consultations with the defendant's expert on discovery-related topics and the forensic production, but to my knowledge, no additional ones were requested.

5. It is certainly true that the circumstances of this case have created an atypical forensic review process. This case involves not only what is, in my experience, an extraordinary volume of data, but also an extraordinary volume of classified data that is particularly sensitive. This required me to conduct the bulk of my work in a specialized review environment suitable for

handling that material and for processing data from a variety of complex proprietary systems. The data pertaining specifically to the theft of data from and otherwise misuse of DEVLAN, and that which is reflected in my analysis and my testimony, is a relatively small portion of an enormous corpus of extremely sensitive material. Overall, I disagree with the Bellovin Affidavit's assertion, in paragraph 13, that "the material provided [to the defense] cannot be utilized to conduct an adversarial forensic examination to put [my] analyses to the test through the scientific method." As part of my work on this case, I participated in the review and preparation of materials that were provided to the defense in discovery. I consulted regularly with the prosecutors handling this case regarding the Government's discovery obligations, and I have been repeatedly made aware of the Government's obligation to produce both the material that I relied on to support my forensic conclusions and any material that would be exculpatory of the defendant. I am certain that the materials that have been provided to the defense are sufficient for a competent digital forensic expert to review and examine the facts to which I testified, and to verify or dispute the conclusions that I drew from those facts, to which I also testified at the trial in this matter in 2020.

6. It is not the case, as asserted in paragraph 11 of the Bellovin Affidavit, that it is not possible to reproduce the analyses and tests necessary to confirm or deny my conclusions without access to the complete mirror images of the ESXi and FS01 servers. While it is true that the most complete forensic image is typically the starting point for digital forensic analysis, not all content in a complete mirror image is relevant to forensic analysis. For example, in this case, the FS01 server (also referred to as the "NetApp"), stores an enormous quantity of classified CIA data, including not only the daily backup files for the Atlassian products used by CIA employees to conduct work on DEVLAN, but also the final copies of cyber tools, user home directories, and other bulk digital storage. Although the forensic artifacts pertaining to the daily backup files were

relevant to my analysis, the overwhelming majority of other files on the NetApp were not relevant. Overall, there were only very few instances in which the content of user-created files, which comprise the bulk of the content of the servers at issue, was relevant to my analysis, such as files that contained cryptographic public and private keys used to authenticate users logging in to various parts of the system, or that would be relevant to any competent forensic examiner reviewing the evidence in this case. My analysis in this case was principally based on system-generated files, such as log files, or file metadata.

7. In addition, as a general matter, the vast majority of my conclusions and testimony in this case was not based on "tests" I ran. On the contrary, and particularly with regard to the defendant's own activities on DEVLAN, my testimony was principally an explanation of the meaning of particular computer commands, filenames, and other terms that appear in log files of activity on various parts of DEVLAN. For example, I testified about a series of records of logged activity in a log file—"viclient-7-0002.log"—found on the defendant's DEVLAN computer. I explained what a "viclient" log file is—meaning the type of log generated by the user, or "client," of vSphere software, which allows a user to connect to different virtual machines and manage those services. I testified as to the times when that activity was recorded, which is evident from the timestamps visible in the logs, and what the technical terms in those log files mean. In particular, I testified that those logs recorded the "client"—that is, the defendant's computer—sending a series of commands on the evening of April 20, 2016, to (i) view the available "snapshots" (meaning copies of the state and data of a virtual machine at the specific time when the snapshot was taken) for the Confluence virtual machine, (ii) create a new snapshot entitled "bkup," (iii) revert the Confluence virtual machine to an earlier snapshot named "bk 4-16-2016," (iv) later re-revert the Confluence virtual machine to the same snapshot entitled "bkup" that had

been created earlier that evening, (v) again view a list of available snapshots, and (vi) delete the snapshot "bkup." My testimony in this regard was not based on any "tests," but rather was based on my training and experience with viclient log files, my knowledge of what those files record, my explanation of what those commands do when executed on a system, and my reading of data evident in those files, such as timestamps and filenames. Another expert with similar expertise would be able to review those log files and verify or dispute my conclusions—it would not be necessary to review the entire mirror image of the computer from which that log file was extracted. This is just one example of the type of expert testimony that I provided in this case. At the trial in this case in 2020, I testified with the aid of an exhibit that collated a variety of forensic artifacts, which was introduced as Government Exhibit 1703. From pages 38 to 173 of that exhibit, you can see that all of the computer forensics that I testified about was the existence of certain lines in log files or unallocated space found on particular computers and the meaning of those terms.

8. I participated in the preparation of a standalone laptop (the "Standalone") containing additional discovery materials that was made available to the defendant's expert in November 2019 at a CIA facility, and which is referred to in paragraph 16 of the Bellovin Affidavit. The Bellovin Affidavit's description of the material contained on that Standalone is, however, misleading. The Standalone contains (i) unredacted copies of the Stash repositories for any CIA tool for which source code had been released by WikiLeaks; (ii) unredacted copies of all Stash documentation that had been released by WikiLeaks; (iii) all Stash commit logs for all projects released by WikiLeaks, redacting only the usernames of the individuals who "committed," or saved, particular versions of those projects; (iv) unredacted copies of the March 2 and 3, 2016 Confluence backups; and (v) unredacted copies of the April 15 and 17, 2016 Crowd authentication SQL databases. The only files on the Standalone for which content was redacted were the Stash

commit logs, and the only redactions were to the usernames. These git repositories and associated commit loges were restored from the July 31, 2016 Stash backup. The process to rebuild the DEVLAN Stash server and the above processes and procedure were documented in my notes labeled "how_to_rebuild_stash." The Standalone is also configured with MySQL server and Git software to facilitate review of the SQL databases and Stash repositories, as well as OpenOffice and Adobe Acrobat to facilitate review of any documents of interest.

9. In preparing this Declaration, I again reviewed the Standalone, which still contains the same material, and which has remained available for the defendant's expert to review. In respect to paragraph 16 of the Bellovin Affidavit, I would expect any competent forensic examiner that was familiar with the evidence and technology involved in this case to have familiarity with basic "git" commands and the Linux manual "man" command, which provides help for the common usage of commands in a Linux environment. While the Internet is a valuable tool, from my experience working in classified and operational environments, providing your own offline documentation is important and sometimes necessary.

10. The Bellovin Affidavit is correct, in paragraph 20, that the existence of a character encoding error in the Linux script used to create the daily backup files of the Confluence virtual machine was a key part of my conclusion that the data released by WikiLeaks in this case came from one of those daily backup files. It was not the only basis for my conclusion, however, which was also premised on visual differences between the leaked data and the data as it appeared in Confluence (indicating that it had not been taken from a live version of the system), the impracticality of other methods of collecting that data, such as scraping each individual page, which would have produced different information than what was disclosed on WikiLeaks, or exporting the entire Confluence virtual machine, which would be an extremely large file. I am

also unaware of any "complex or technical reason," as referenced but not described in paragraph 34 of the Bellovin Affidavit, why a competent forensic expert would be unable to identify the same error in the backup script that I was. I testified specifically about the portion of the script that contained the error—the "mysqldump" command, which is designed to export the Structured Query Language ("SQL") database used to define the relationships between different tables of data contained in Confluence. The command was misformatted with respect to assigning the proper character encoding during the creation of the backups. I also testified about portions of the SQL databases for the actual Confluence backup files that showed the propagation of that error. It certainly is not the case that complete mirror images of the ESXi Server and the NetApp are necessary, or would even be relevant, to testing that conclusion. The backup script error pertained to the SQL database, so only the SQL database and not the large volume of content files, would be relevant to examining the error. In addition, the backup script was stored in the Confluence virtual machine itself on the ESXi server, and functioned only to backup the Atlassian products running on DEVLAN. The large volume of files on the ESXi and NetApp servers that do not pertain to the Atlassian products would have no relevance to an evaluation of the error in the backup script.

11. The unredacted portion of the Bellovin Affidavit, in paragraphs 23 and 24, discusses the fact that access times for files can be affected by different types of commands. The unredacted portion of the Bellovin Affidavit does not address what information the defendant's expert believes would reveal that the "date accessed" field for the stolen Confluence backup files was modified by a user action other than copying the files, so I am not in a position to respond to any such arguments. Historical information about actions modifying the "date accessed" fields for files stored on the NetApp server would be reflected, if at all, in the NetApp log files. From my

review of those log files, the relevant files retained data for a relatively short period of time and information from April 2016 is not available. The log files that were available from the NetApp server do not include any evidence that the March 3, 2016 Confluence backup files were modified by any user action other than copying the files, and the log files that were available were produced in discovery.

12. The Bellovin Affidavit is correct, in paragraph 31, that the Government has produced a large volume of log files and other information, including unallocated space, to the defense in this case. It is also true that as part of my role in reviewing the evidence collected in this case, I had access to additional information. But any information that I identified as relevant either to the conclusions I drew or to the allegations at issue in this case was produced to the defense. As described above, I reviewed the data that was available to me in conjunction with numerous discussions with the Government prosecutors in this case about the Government's discovery obligations, and I identified anything that I considered to be relevant for production to the defense.

13. The Bellovin Affidavit is correct, in paragraph 36, that I "testified about Mr. Schulte's encrypted private SSH key," but it is false in asserting, in paragraph 36, both that I "did not verify that Mr. Schulte's private key corresponded to the public key file," and that the defendant's expert "lacked access to machine-readable and processable copies of the files purported to be Mr. Schulte's private and public SSH keys." An SSH key pair consists of public and private keys made up of strings of alphanumeric characters that are mathematically generated using a technique referred to as asymmetric cryptography, and can be analyzed using standard cryptographic algorithms to generate a "key fingerprint." I testified at the trial in this case in 2020 that "a key fingerprint is a way to identify—it's a unique identifier for a specific public-private

key pair, and we know this to be the defendant's public-private key that's on his Ubuntu virtual machine . . . [because] I've calculated the fingerprint and compared it, and it matches." (Trial Transcript 990:23-25; 991:1-4). Moreover, the defendant and his expert have access to those keys, which were introduced as Government Exhibits at the previous trial. The defendant has copies of his public key from two locations, from the ESXi server, which was introduced as Government Exhibit 1209-16, and from the April 16, 2016 backup of the Confluence virtual machine, which was introduced as Government Exhibit 1207-7, both of which show that the key was assigned to user "schuljo@devlan.net". The defendant has a copy of his private key from his DEVLAN computer, which was introduced as Government Exhibit 1203-9. The defendant was also provided with the password used to encrypt his private key, which was "KingJosh3000." Any competent expert would be able to calculate the key fingerprint and validate the authenticity of the keypair.

14. I declare under penalty of perjury that, to the best of my knowledge, the foregoing information is true and correct.

Dated: Fairfax County, Virginia
April 29, 2022

Patrick Leedom